**Objective: I do statistic analysis to compare the consumption of each products (Fresh, Milk, Grocery, Frozen, Detergents\_Paper, Delicassen) among all 3 regions so the client tracks their sale.**

**Hypothesis test**

**I apply hypothesis test to each product to compare the average among 3 regions. From here clients understand the consumptions of these products are different between each region (1,2,3) or not.**

**PROC** **IMPORT** OUT=WORK.PROJECT\_FLORATRAN

DATAFILE='\\ssomfs2\shome\ttran23\Documents\ISOM631\Wholesalecustomersdata.xlsx'

DBMS=XLSX REPLACE;

SHEET='Wholesalecustomersdata';

GETNAMES=yes;

**RUN**;

1. **Fresh product**

/\*H0: Means from the consumption of fresh product of all regions are the same.\*/

/\*H1: At least one of the mean is different from the rest.\*/

**PROC** **ANOVA** DATA=WORK.PROJECT\_FLORATRAN;

class Region;

model Fresh = Region;

means Region / hovtest welch;

**RUN**;

/\* We fail to reject H0 since the P-Value is equal to 0.4330 > 0.05.

All means from fresh product of all regions are the same \*/

1. **Milk product**

/\*H0: Means from the consumption of Milk product of all regions are the same.\*/

/\*H1: At least one of the mean is different from the rest.\*/

**PROC** **ANOVA** DATA=WORK.PROJECT\_FLORATRAN;

class Region;

model Milk = Region;

means Region / hovtest welch;

**RUN**;

/\* We fail to reject H0 since the P-Value is equal to 0.5910 > 0.05.

All means from milk product of all regions are the same \*/

1. **Grocery product**

/\*H0: Means from the consumption of Grocery product of all regions are the same.\*/

/\*H1: At least one of the mean is different from the rest.\*/

**PROC** **ANOVA** DATA=WORK.PROJECT\_FLORATRAN;

class Region;

model Grocery = Region;

means Region / hovtest welch;

**RUN**;

/\* We fail to reject H0 since the P-Value is equal to 0.8569 > 0.05.

All means from grocery product of all regions are the same \*/

1. **Frozen product**

/\*H0: Means from the consumption of frozen product of all regions are the same.\*/

/\*H1: At least one of the mean is different from the rest.\*/

**PROC** **ANOVA** DATA=WORK.PROJECT\_FLORATRAN;

class Region;

model Frozen = Region;

means Region / hovtest welch;

**RUN**;

/\* We reject H0 since the P-Value is equal to 0.0429 < 0.05.

All means from frozen product of all regions are not the same \*/

1. **Detergents\_Paper product**

/\*H0: Means from the consumption of Detergents\_Paper product of all regions are the same.\*/

/\*H1: At least one of the mean is different from the rest.\*/

**PROC** **ANOVA** DATA=WORK.PROJECT\_FLORATRAN;

class Region;

model Detergents\_Paper = Region;

means Region / hovtest welch;

**RUN**;

/\* We fail to reject H0 since the P-Value is equal to 0.3920 > 0.05.

All means from Detergents\_Paper product of all regions are the same \*/

1. **Delicassen**

/\*H0: Means from the consumption of Delicaseen product of all regions are the same.\*/

/\*H1: At least one of the mean is different from the rest.\*/

**PROC** **ANOVA** DATA=WORK.PROJECT\_FLORATRAN;

class Region;

model Delicassen = Region;

means Region / hovtest welch;

**RUN**;

/\* We fail to reject H0 since the P-Value is equal to 0.7193 > 0.05.

All means from Delicassen product of all regions are the same \*/

**Histogram Charts**

**I create histogram on all products**

/\*Creating Frequency Distributions\*/

**PROC** **FREQ** DATA=WORK.PROJECT\_FLORATRAN;

TABLE Fresh Milk Grocery Frozen Detergents\_Paper Delicassen ;

**RUN**;

/\*Creating Histogram\*/

**PROC** **UNIVARIATE** DATA=WORK.PROJECT\_FLORATRAN;

Histogram Fresh / midpoints= **200** to **85000** by **25000**;

Histogram Milk / midpoints= **100** to **85000** by **25000**;

Histogram Grocery / midpoints= **100** to **85000** by **30000**;

Histogram Frozen / midpoints= **100** to **75000** by **25000**;

Histogram Detergents\_Paper / midpoints= **100** to **75000** by **25000**;

Histogram Delicassen / midpoints= **100** to **75000** by **25000**;

**Run**;

/\* Average of all products in 3 regions \*/

**PROC** **MEANS** N MEAN MIN MAX DATA=Work.PROJECT\_FLORATRAN;

class region;

var Fresh Milk Grocery Frozen Detergents\_Paper Delicassen;

**RUN**;

**Regression or Forecasting models**

**I set regression analysis to find the best model that explains the relationship between each variable so the client can predict future price they should effectively set new price on each product. Also, clients can base on this to know when the sale market is high to announce a discount offer or make their business strategy to boost their sale.**

**/\* Predict Fresh price \*/**

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

model Fresh = Milk Grocery Frozen Detergents\_Paper Delicassen / vif;

**RUN**;

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

Model Fresh = Milk Frozen Detergents\_Paper Delicassen / SELECTION=Backwards SLS=**0.05** ADJRSQ;

**RUN**;

/\* Y = (M \* X )+ b = (0.32 \* Milk) + (0.77 \* Frozen) + (-0.495 \* Detergent) + 9182.35

**/\* Predict Milk price \*/**

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

model Milk = Fresh Grocery Frozen Detergents\_Paper Delicassen / vif;

**RUN**;

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

Model Milk = Fresh Frozen Delicassen / SELECTION=Backwards SLS=**0.05** ADJRSQ;

**RUN**;

/\* Y = (M \* X )+ b = (1.06349 \* Delicassen) + 4174.58306

**/\* Grocery \*/**

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

model Grocery = Milk Fresh Frozen Detergents\_Paper Delicassen / vif;

**RUN**;

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

Model Grocery = Milk Fresh Frozen Detergents\_Paper Delicassen / SELECTION=Backwards SLS=**0.05** ADJRSQ;

**RUN**;

/\* Y = (M \* X )+ b = (0.18569 \* Milk) + (0.02886 \* Fresh) + (1.64948 \* Detergent) + (0.27010 \* Delicassen) + 1363.8 \*/

**/\*Frozen\*/**

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

model Frozen = Milk Grocery Fresh Detergents\_Paper Delicassen / vif;

**RUN**;

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

Model Frozen = Milk Fresh Delicassen / SELECTION=Backwards SLS=**0.05** ADJRSQ;

**RUN**;

/\* Y = (M \* X )+ b =(0.10216 \* Fresh) + (0.56089 \* Delicassen) + 990.65538

**/\* Detergents\_Paper \*/**

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

model Detergents\_Paper = Fresh Milk Grocery Frozen Delicassen / vif;

**RUN**;

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

Model Detergents\_Paper = Fresh Milk Grocery Frozen Delicassen / SELECTION=Backwards SLS=**0.05** ADJRSQ;

**RUN**;

/\* Y = (M \* X )+ b = (-0.02128 \* Fresh) + (0.04029 \* Milk) + (0.45147 \* Grocery) + (-0.03925 \* Frozen) + (-0.18857 \* Delicassen) -278.36955

**/\* Delicassen \*/**

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

model Delicassen = Fresh Milk Grocery Frozen Detergents\_Paper / vif;

**RUN**;

**PROC** **REG** data=WORK.PROJECT\_FLORATRAN;

Model Delicassen = Fresh Milk Frozen Detergents\_Paper / SELECTION=Backwards SLS=**0.05** ADJRSQ;

**RUN**;

/\* Y = (M \* X )+ b = (0.20308 \* Milk) + (0.17019 \* Frozen) + (-0.14427 \* Detergent) + 240.68279 \*/